



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### **Multimomics analysis provides insight into the laboratory evolution of *Escherichia coli* towards the metabolic usage of fluorinated indoles**

**Citation for published version:**

Agostini, F, Sinn, L, Petras, D, Schipp, CJ, Kubyshkin, V, Berger, AA, Dorrestein, PC, Rappsilber, J, Budisa, N & Kokschi, B 2020, 'Multimomics analysis provides insight into the laboratory evolution of *Escherichia coli* towards the metabolic usage of fluorinated indoles', *ACS Central Science*.  
<https://doi.org/10.1021/acscentsci.0c00679>

**Digital Object Identifier (DOI):**

[10.1021/acscentsci.0c00679](https://doi.org/10.1021/acscentsci.0c00679)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

ACS Central Science

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Multimomics Analysis Provides Insight into the Laboratory Evolution of *Escherichia coli* toward the Metabolic Usage of Fluorinated Indoles

Federica Agostini, Ludwig Sinn, Daniel Petras, Christian J. Schipp, Vladimir Kubyshkin, Allison Ann Berger, Pieter C. Dorrestein, Juri Rappsilber, Nediljko Budisa,\* and Beate Kokschi\*



Cite This: <https://dx.doi.org/10.1021/acscentsci.0c00679>



Read Online

ACCESS |



Metrics & More

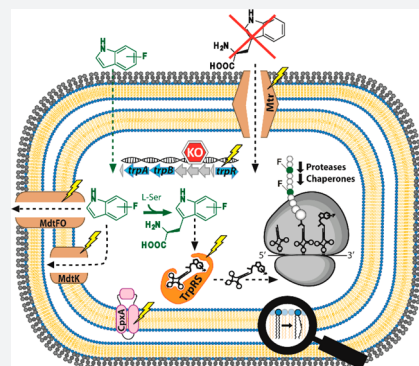


Article Recommendations



Supporting Information

**ABSTRACT:** Organofluorine compounds are known to be toxic to a broad variety of living beings in different habitats, and chemical fluorination has been historically exploited by mankind for the development of therapeutic drugs or agricultural pesticides. On the other hand, several studies so far have demonstrated that, under appropriate conditions, living systems (in particular bacteria) can tolerate the presence of fluorinated molecules (e.g., amino acids analogues) within their metabolism and even repurpose them as alternative building blocks for the synthesis of cellular macromolecules such as proteins. Understanding the molecular mechanism behind these phenomena would greatly advance approaches to the biotechnological synthesis of recombinant proteins and peptide drugs. However, information about the metabolic effects of long-term exposure of living cells to fluorinated amino acids remains scarce. Hereby, we report the long-term propagation of *Escherichia coli* (*E. coli*) in an artificially fluorinated habitat that yielded two strains naturally adapted to live on fluorinated amino acids. In particular, we applied selective pressure to force a tryptophan (Trp)-auxotrophic strain to use either 4- or 5-fluorotryptophan as essential precursors for the *in situ* synthesis of Trp analogues, followed by their incorporation in the cellular proteome. We found that full adaptation to both fluorinated Trp analogues requires a low number of genetic mutations but is accompanied by large rearrangements in regulatory networks, membrane integrity, and quality control of protein folding. These findings highlight the cellular mechanisms behind the adaptation to unnatural amino acids and provide the molecular foundation for bioengineering of novel microbial strains for synthetic biology and biotechnology.

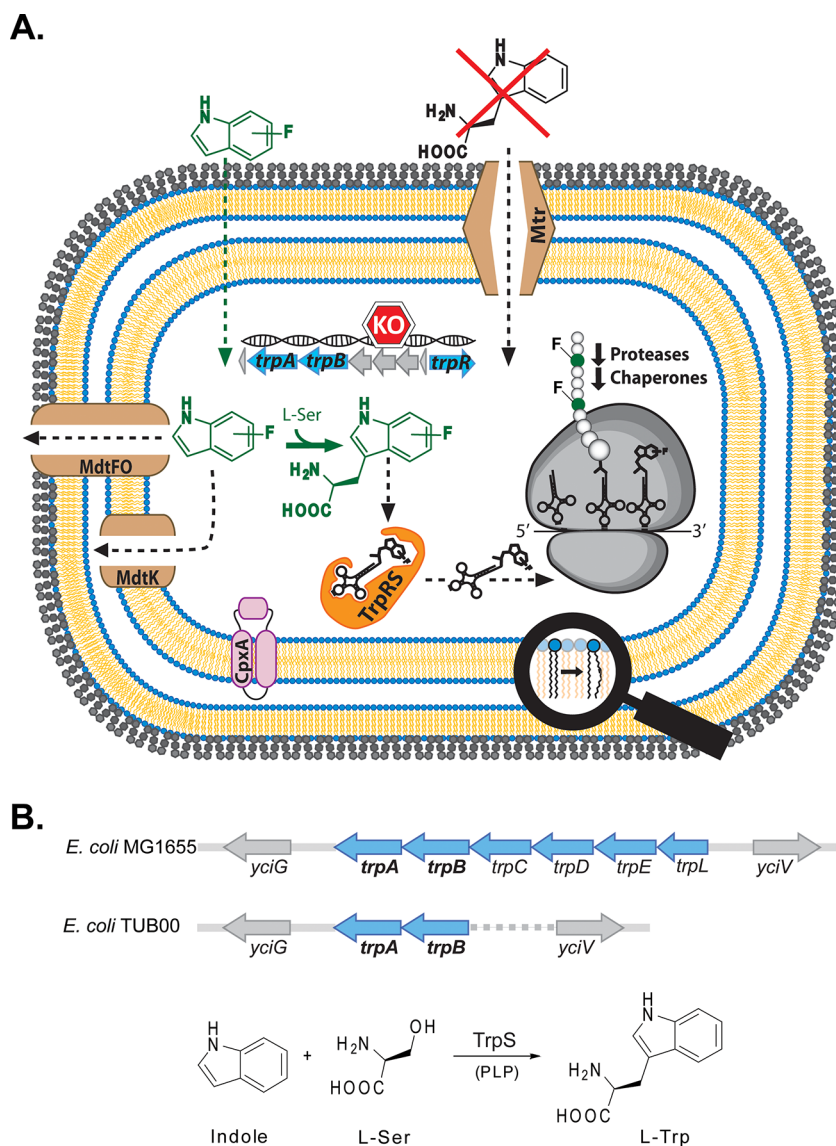


For billions of years, living organisms have used mainly six chemical elements (carbon, hydrogen, nitrogen, oxygen, phosphorus, and sulfur) out of the 118 available on Earth for the synthesis of the core macromolecules of life (DNA, RNA, proteins, lipids, and carbohydrates). Fluorine, despite being used by humans for various innovative syntheses, has been neglected by evolution and is found only in very few natural molecules.<sup>1</sup> The factors that make fluorine poorly suitable for biochemical reaction are its very low bioavailability of fluoride ions ( $F^-$ ) in the oceans (1.3 ppm), the high heat of hydration (ca. 120 kcal mol<sup>-1</sup>), and the least tendency to be oxidized by haloperoxidases among all halogens.<sup>2</sup> The high strength and polarization of the C–F bond alter the geometry, conformation, and interactions of molecules; hence, its incorporation has proven to be an extremely powerful method to modulate stability and/or activity of a vast variety of materials, fine chemicals, drugs, and pesticides.<sup>3–5</sup> Nowadays, molecular biological techniques enable scientists to artificially label complex biological macromolecules such as peptides and proteins with fluorine via incorporation of fluorinated amino acids.<sup>6</sup> The main advantage of introducing fluorinated domains is that they can drive the processes of protein–protein

interaction and folding<sup>7</sup> or tune resistance against proteolytic degradation.<sup>8</sup> However, the properties that make fluorine attractive for modulating protein properties *in vitro* often represent a threat *in vivo*. Indeed, once introduced in a cellular environment, many fluorine-containing molecules behave as metabolic stressors.<sup>9</sup>

As early as the 1960s, it was observed that fluorinated analogues of standard (“canonical”) amino acids could be fed to bacterial cultures leading to their partial incorporation into endogenous proteins.<sup>10,11</sup> However, cell growth was inhibited.<sup>12</sup> Later on, engineering efforts resulted in the creation of laboratory strains able to grow on fluorotryptophan analogues<sup>13,14</sup> and trifluoroleucine.<sup>15</sup> The whole field was brought forward by the experimental effort of Wong<sup>13</sup> which

Received: May 27, 2020



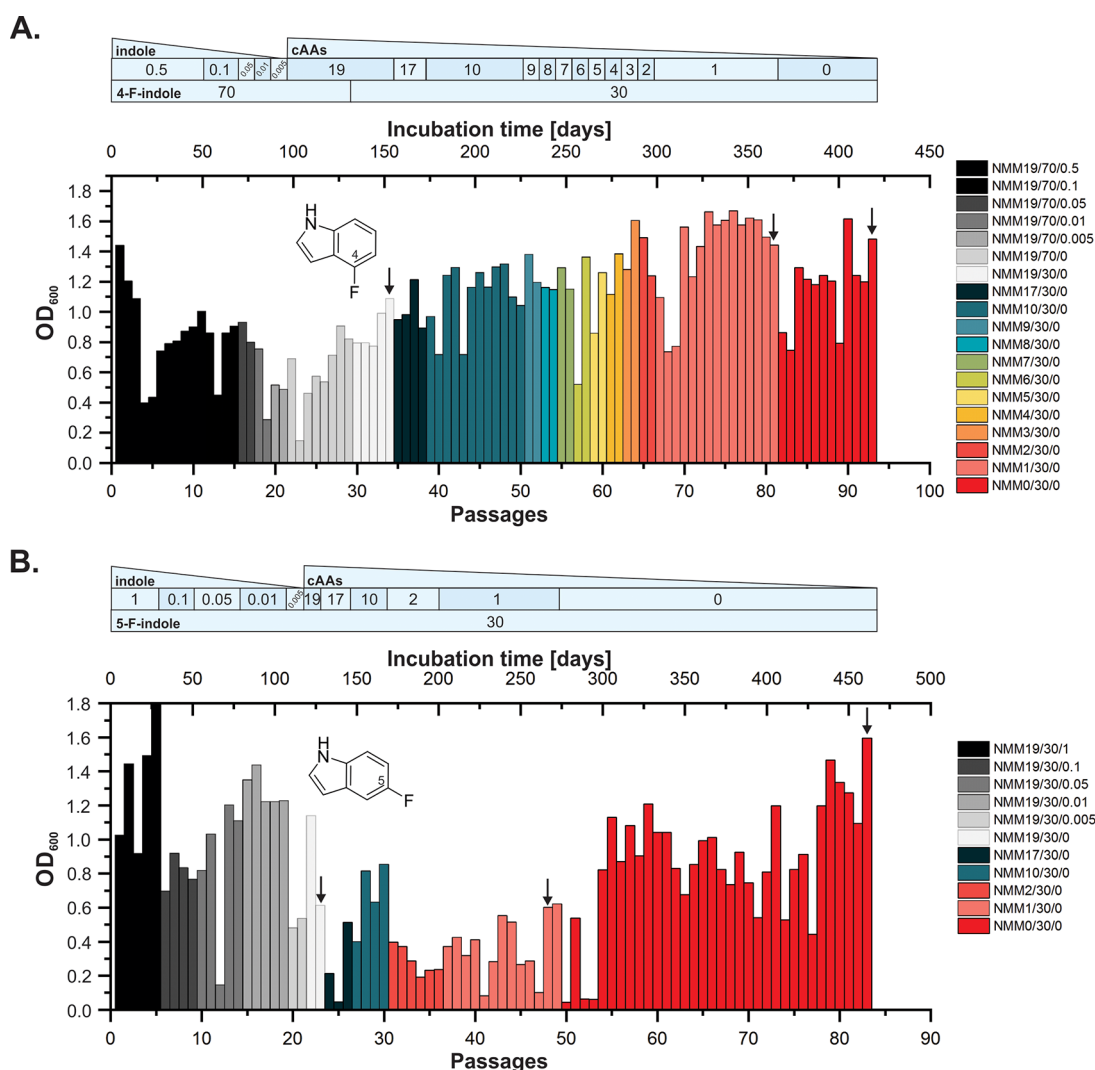
**Figure 1.** (A) ALE experimental setup and overview of the changes observed in the *E. coli* strains adapted to fluorindoles that marked a divergence from TUB00. Abbreviations include the following: *trpA* (Trp synthase  $\alpha$  subunit); *trpB* (Trp synthase  $\beta$  subunit); *trpR* (*trp* operon transcriptional repressor); TrpRS (tryptophanyl-tRNA synthetase); Mtr (Trp permease); MdtF, MdtO, and MdtK (multidrug efflux pumps); and CpxA (sensor histidine kinase). Under natural conditions, cells can uptake extracellular Trp. However, during the ALE cultivation, Trp was not supplied; hence, its structure is crossed out with a red “X”. (B) *trp* operon of *E. coli* in the parent strain MG1655 and in the Trp-auxotrophic derivative TUB00 (used for ALE cultivation) after deletion of the genes *trpLEDC*. The reaction catalyzed by TrpS (which requires pyridoxal-5'-phosphate, PLP, as cofactor) is shown below.

replaced all tryptophan (Trp) residues in the proteome of a *Bacillus subtilis* mutant by 4-fluorotryptophan. The strain was treated with *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine to promote mutagenesis. Almost 20 years later, Bacher and Ellington used a serial-dilution-growth approach to fully adapt *E. coli* to 4-fluorotryptophan.<sup>14</sup> Their work proved the high-level exchange of Trp by 4-fluorotryptophan in bacterial soluble proteins. The natural counterpart Trp, however, was continuously present in cultures (80 nM) as commercial preparations of fluorinated analogues contained traces of Trp. Recently, Chan, Wong, et al. re-evaluated the classical Wong experiment and included 5-fluorotryptophan and 6-fluorotryptophan in their studies.<sup>16,17</sup>

Genomic and transcriptomic analyses of tolerant strains<sup>16–18</sup> led to the postulation of the “oligogenic barrier” hypothesis, which states that the main limitation to the proteomic

replacement of canonical amino acids lies in a small number of proteins that malfunction upon incorporation of fluorinated analogues. However, the observations of the different studies do not converge to a common theory. Hence, to date it has not been clarified which molecular mechanisms need to be targeted in order to enable the growth of cells on fluorinated amino acids.

In this study, we intended to investigate the biochemical rearrangement that living systems must undergo in order to accommodate fluorinated amino acids and proteins in their metabolism. The aim of the present work was to systematically monitor the metabolic effects of the long-term presence of fluorine in the protein chemical repertoire of living cells and to define whether different fluorinated amino acids activate common regulatory systems and networks. To this end, we carried out two parallel adaptive laboratory evolution (ALE)



**Figure 2.** Cultivation scheme of *E. coli* ALE toward usage of (A) 4-fluoroindole; (B) 5-fluoroindole as precursors for the synthesis of Trp analogues. Optical density ( $OD_{600}$ ) at the reinoculation step (“passage”) is plotted against days of incubation and number of passages. The color code refers to the composition of minimal medium (NMM, Tables S3 and S4), NMMa/b/c where *a* is the number of amino acids supplied, *b* the concentration of fluorindole, and *c* the concentration of indole, both  $\mu$ M. Black arrows indicate the isolates selected for multiomics analysis, referred to as early (4TUB34 and 5TUB23), intermediate (4TUB81 and 5TUB48), and final time points (4TUB93 and 5TUB83).

experiments that coupled for the first time alternative protein synthesis in *Escherichia coli* (*E. coli*) with *in situ* synthesis of fluorinated amino acid analogues via enzymatic conversion of commercially available fluorinated precursors. In particular, selective pressure was exerted on a Trp-auxotrophic strain which, in order to grow, was forced to convert 4- and 5-fluoroindole into 4- and 5-fluorotryptophan (respectively) via a single-step reaction catalyzed by the endogenous enzyme Trp synthase (TrpS). Subsequently, the newly synthesized analogues must be used by the tryptophanyl-tRNA synthetase for the acylation of  $tRNA_{CCA}$  and ribosomal protein synthesis. Trp is an exceptionally attractive target for protein engineering because it is encoded by a single codon (UGG) in the genetic code, and there are no salvage pathways for its biosynthesis in *E. coli*. As also mentioned above, the replacement of Trp by 4-fluorotryptophan into the proteome of *E. coli* via long-term cultivation had been demonstrated before by Bacher and Ellington.<sup>14</sup> However, the authors reported that the fluorinated amino acid acted as a toxic metabolite for the cells, and they

eventually achieved only partial incorporation, while the viability of their strain decreased over time.

In contrast, we report here the selection of clones that gained the ability to live under progressively more stringent conditions, characterized by the absence of Trp and all other canonical amino acids supplemented in the cultivation medium. We adapted a procedure<sup>19</sup> that previously enabled the trophic replacement of Trp by a sulfurated analogue by feeding the corresponding indole precursor to an *E. coli* strain lacking the *trp* operon. In our ALE experiments, we generated a new Trp auxotrophic strain which we then cultivated in synthetic minimal medium supplemented with a mixture of indole and either 4- or 5-fluoroindole. Over the time span of several months, we applied increasing selection pressure by decreasing the availability of the Trp precursor (indole) until complete depletion. This method enabled the natural adaptation of *E. coli* to two different fluorinated indole/Trp analogues and thus proved efficient for the generation of cells with a newly adapted metabolism. It must be noted that indole and its fluorinated analogues are not less toxic than the



corresponding amino acids derivatives. In fact, they exert a concentration-dependent inhibition of cell division by altering the electrostatic potential of cellular membranes.<sup>20</sup> Nevertheless, our long-term adaptation setup enabled cells to repurpose these toxic metabolites into substrates for the synthesis of new translationally active amino acids.

In order to investigate whether the adaptation process had involved the simultaneous rearrangement of more than one class of biomolecules, isolates from different time points of the ALE experiment were systematically studied by integrating genomic, proteomic, and metabolomic analyses. An adaptation model is proposed based on these data, highlighting that stress response, quality of protein folding, and membrane integrity are the most critical factors influencing the creation of fluorinated proteomes.

## RESULTS

Based on literature reports and our own determination of the catalytic parameters of TrpRS, we knew that 4- and 5-fluorotryptophan would be suitable substrates for proteome synthesis (see Table S1). In preliminary experiments, we had observed that 4- and 5-fluoroindole were almost isosteric to indole but exhibited higher lipophilicity and polarity (Figure S1). The indole/Trp background and stable Trp-auxotrophy were installed by deleting the genes of both the biosynthetic (*trpLEDC*; *trp* operon) and degradation (*tnaA*; encoding the enzyme tryptophanase) pathways. The new strain was named TUB00 ( $\Delta trpLEDC$ ,  $\Delta tnaA$ , Figure 1 and Figure S2). Resistance to Trp was not achieved because the strains naturally retained the copies of the genes encoding TrpS (tryptophanyl synthase, *trpBA*) and TrpRS (tryptophanyl-tRNA synthetase, *trpS*) which are able to recognize, respectively, indole and Trp as well as their analogues with similar efficiency.

The ALE setup included continuous cultivation in synthetic minimal medium (NMM) in shake flasks and serial reinoculation of parallel cultures at the late exponential phase, to enable full assimilation of fluorindoles (see Method 6 in the SI). Thereby, the concentration of the major growth nutrient indole was gradually decreased from 0.5  $\mu$ M (in the case of the 4-fluoroindole experiments), or 1  $\mu$ M (in the case of the 5-fluoroindole experiments), to zero, while keeping the levels of respective fluorindoles and other nutrients (in particular glucose, ammonium, and phosphate) constant (Tables S3 and S4). Contamination of 4- and 5-fluoroindole by indole was not detectable by GC-MS analysis (Figures S3–S8, Method 2 in the SI).

The ALE cultivation for each indole analogue was started with three biological replicate cultures of the ancestor strain, which evolved into separated lineages (see the SI, Method 6, for experimental details). Of three lineages, only one survived the conditions imposed (i.e., gradual removal of amino acids from the cultivation medium). The cells remained viable for over 1 year of cultivation and eventually acquired the ability to grow in NMM0 (i.e., containing no amino acids) supplied exclusively with 4- or 5-fluoroindole. This demonstrated that *E. coli* had developed the ability to use the fluorinated substrates as metabolic intermediates, i.e., “xeno-nutrients”. Overall, we propagated parallel serial cultures for 825 generations and 93 serial reinoculation steps in the case of 4-fluoroindole ALE (Figure 2A) and 678 generations and 83 serial reinoculation steps in the case of 5-fluoroindole ALE (Figure 2B). The strains adapted to 4-fluoroindole are referred to below as

“4TUBX” and those adapted to 5-fluoroindole as “5TUBX”, where X corresponds to the respective serial reinoculation step number (“passage”). Within each lineage, 3 time points of the ALE cultivation were selected, and for each of them, 3 colonies were used as biological replicates for the multiomics analysis. At the end of the ALE, the adapted strains had become facultative fluorotryptophan/Trp users.

The capability of the final isolates 4TUB93 and 5TUB83 to produce fully fluorinated proteins was tested by expressing the enhanced green fluorescent protein EGFP, which has one single Trp residue at position 57 in its primary sequence. The cells were transformed with a plasmid carrying the sequence of C-terminal six-histidine-tagged EGFP (EGFP-H6); the protein was expressed at 30 °C in NMM0 containing either 4- or 5-fluoroindole and subsequently purified via Ni-NTA affinity chromatography (see the SI, Method 7). The identity of the purified product was confirmed by LC-MS.

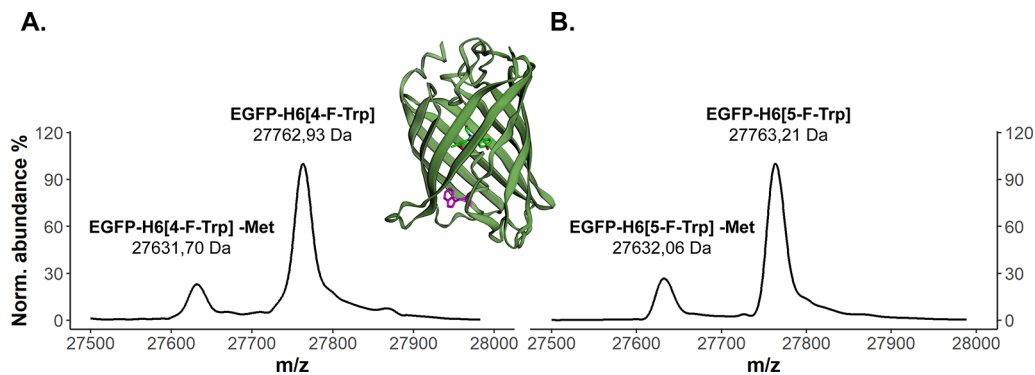
The observed masses matched the theoretical values expected for EGFP-H6 variants labeled with 4- and 5-fluorotryptophan, after maturation of the chromophore. In both experiments, the main mass corresponded to the full-size protein, and a second mass was assigned to the protein after cleavage of the initial methionine (Table 1 and Figure 3). The

**Table 1. Theoretical and Observed Mass Values of His-Tagged Enhanced Green Fluorescent Protein (EGFP-H6) after Expression in 4TUB93 and 5TUB83 (The Values Correspond to the Protein after Chromophore Maturation)**

	theoretical mass	observed mass	
		4TUB93	5TUB83
EGFP-H6 [F-Trp]	27 763.33 Da	27 762.93 Da	27 763.21 Da
EGFP-H6 [F-Trp]–Met	27 632.14 Da	27 631.70 Da	27 632.06 Da

successful expression of EGFP-H6 [F-Trp] demonstrated that, in the adapted strains, Trp had been completely replaced by 4- and 5-fluorotryptophan during protein synthesis.

Subsequently, we proceeded to characterize whole-cell isolates by means of genome sequencing, quantitative proteomic, and nontargeted metabolomic analyses. First, genomic mutations that accumulated during the two parallel ALE experiments were identified by whole-genome sequencing of isolates of 4TUB and 5TUB at early, intermediate, and final time points by means of an Illumina HiSeq 4000 sequencing platform (Table S6 and supplementary data; sequencing and evaluation were performed by BGI Hong Kong). The impact of mutations on structure and function of the corresponding protein product was predicted by PROVEAN v1.1 (Protein Variation Effect Analyzer, Table S7).<sup>21,22</sup> Notably, *E. coli* adapted to whole-proteome fluorination by a relatively low number of mutations. Of more than 20 000 TGG (Trp) codons present in the genome of *E. coli*, only one was mutated during 5-fluoroindole ALE (in the gene *mrr*, encoding methylated adenine and cytosine restriction protein, Trp105Stop). This suggests a remarkable tolerance of the bacterial proteome toward global incorporation of fluorotryptophans. No major changes of ribosomal proteins were required to catalyze protein synthesis with 4- and 5-fluorotryptophan, which is in agreement with the well-known tolerance of the ribosome toward different amino acid analogues.<sup>23</sup> Only during 4-fluoroindole ALE did single point



**Figure 3.** Deconvoluted mass spectra of His-tagged enhanced green fluorescent protein (EGFP-H6) expressed in (A) 4TUB93 and (B) 5TUB83. Structure of the enhanced green fluorescent protein from *Aequorea victoria* (EGFP, structure 2Y0G deposited in PDB) figure generated with the EzMol interface.<sup>21</sup> Trp57 is highlighted in pink.

**Table 2. Genomic Mutations in Genes Involved in Transcription as Well as in Indole, Trp, and Protein Metabolism Detected at Early, Intermediate, and Final Time Points of 4-Fluoroindole ALE (4TUB34, 4TUB81, and 4TUB93) and of 5-Fluoroindole ALE (5TUB23, 5TUB48, and 5TUB83) [Deletion of a Nucleotide in *ptsG* Induces a Translational Frameshift (FS)]**

gene	protein	ALE isolates					
		4TUB34	4TUB81	4TUB93	5TUB23	5TUB48	5TUB83
<i>trpS</i>	tryptophanyl-tRNA synthetase	Glu15Asp	Glu15Asp	Glu15Asp			Met187Leu
<i>trpR</i>	<i>trp</i> operon transcriptional repressor				Arg84Cys	Arg84Cys	Arg84Cys
<i>ptsG</i>	integral membrane permease IICB <sup>Glc</sup>				Leu250Arg	Ile25-FS	Ile25-FS
<i>mtr</i>	Trp permease				Ser94Pro	Ser94Pro	Ser94Pro
<i>mdtF</i>	multidrug resistance protein MdtF	Asp759Gly	Asp759Gly	Asp759Gly			
<i>mdtK</i>	multidrug resistance protein MdtK						Gln305Stop
<i>mdtO</i>	multidrug resistance protein MdtO						Ile104Ser
<i>rpoA</i>	RNA polymerase subunit $\alpha$		Val264Ala	Val264Ala			
<i>rpoC</i>	RNA polymerase subunit $\beta'$						Asp1208Ala
<i>rpsA</i>	30S ribosomal protein S1		Lys158Gln	Lys158Gln		Gln355Pro	
<i>rpsJ</i>	30S ribosomal protein S10		Thr28Ala	Thr28Ala			
<i>ileS</i>	isoleucyl-tRNA synthetase						Tyr727Ser
<i>glyQ</i>	glycyl-tRNA synthetase $\alpha$ subunit		Glu48Ala	Glu48Ala			
<i>ptrA</i>	protease III						Phe827Cys
<i>ftsH</i>	ATP-dependent metalloprotease FtsH					Leu563Arg	Leu563Arg
<i>cpxA</i>	sensor histidine kinase CpxA		Glu355Gly	Glu355Gly			Tyr364Ser

mutations occur at the level of the 30S ribosomal proteins S1 (*rpsA*, Table 2) and S10 (*rpsJ*, Table 2). Mutations of other genes involved in protein biosynthesis included isoleucyl-tRNA synthetase (*ileS*, Table 2) in 5-fluoroindole ALE and glycyl-tRNA synthetase (*glyQ*) in 4-fluoroindole ALE (this latter categorized by PROVEAN as neutral, Table S7). Two genes encoding proteases (*ptrA* and *ftsH*, Table 2 and Table S7) were mutated over the course of 5-fluoroindole ALE, which might increase the cellular tolerance toward fluorinated proteomes. Point mutations were found in the genes encoding two subunits of the RNA polymerase, *rpoA* and *rpoC* (Table 2) for 4- and 5-fluoroindole ALE, respectively. Single nucleotide polymorphisms of members of this enzymatic complex have been observed also in other ALE studies<sup>24</sup> and might help to reconfigure transcription under different environmental conditions. Notably, Yu et al. report mutation of *rpoC* in their 4-fluorotryptophan adapted strain; however, that mutation targeted a different residue from the one in our study.

The gene *trpS*, encoding TrpRS, was mutated in both ALE experiments. We reasoned that the mutant enzymes might improve the usage of fluorinated Trp analogues for protein synthesis, especially in the case of 5-fluorotryptophan, where mutation of *trpS* is observed at the end of ALE. The mutation

of *trpS* in 4-fluoroindole ALE was categorized as neutral (Table S7), which is consistent with our preliminary observation that 4-fluorotryptophan is relatively efficiently activated by TrpRS (Table S1); hence, adaptation did not require substantial modification of the enzyme.

Notably, during 5-fluoroindole ALE, a mutation in *trpR*, encoding the transcriptional repressor TrpR of the *trp* operon, altered the conserved residue Arg84, which is critical for binding to DNA via electrostatic interactions (Table 2, Table S7 and Figure S13).<sup>25</sup> TrpR regulates the catabolite-repression of Trp biosynthesis and downregulates the expression of *trpA* and *trpB*, encoding the Trp synthase. Replacement of the positively charged guanidinium group of Arg84 by the uncharged side-chain of Cys is expected to decrease the efficiency of TrpR as transcriptional repressor and to upregulate Trp biosynthesis. During the adaptation toward 4-fluoroindole, wild type TrpR is conserved, and TrpS is downregulated (see Table 2 and proteomics Auxiliary Table 2). Another mutation possibly involved in Trp biosynthesis is the frameshift mutation of the membrane glucose transporter PtsG/IICB<sup>Glc</sup> (*ptsG*, Table 2) that appeared at the intermediate time point of 5-fluoroindole ALE. PtsG/IICB<sup>Glc</sup> is downregulated in all ALE isolates (see Table 2 and

**Table 3. Differential Abundance of Protein Chaperones, Proteases, and Members of the CpxAR-System Regulon at Early and Final Time Points of 4-Fluoroindole ALE (4TUB34 and 4TUB93, Respectively) and 5-Fluoroindole ALE (5TUB23 and 5TUB83, Respectively), Compared to TUB00<sup>a</sup>**

Gene	Protein	Change of abundance in ALE isolate vs TUB00							
		4TUB34		4TUB93		5TUB23		5TUB83	
		log2-fold	-Log (p-value)	log2-fold	-Log (p-value)	log2-fold	-Log (p-value)	log2-fold	-Log (p-value)
<i>ppiA</i>	Peptidyl-prolyl <i>cis-trans</i> isomerase A	1.883	3.739	0.032	0.056	-2.939	2.759	0.766	1.655
<i>fkpA</i>	Peptidyl-prolyl <i>cis-trans</i> isomerase FkpA	1.738	3.546	0.445	1.035	-0.330	1.173	0.755	2.561
<i>hscB</i>	Co-chaperone protein HscB	1.246	3.517	0.302	0.915	0.658	3.789	0.598	1.386
<i>spy</i>	Periplasmic chaperone Spy	-1.596	3.376	-3.312	4.506	-1.103	2.698	-2.530	3.235
<i>hchA</i>	Protein deglycase I	-1.007	3.160	-3.314	4.178	-0.378	0.485	-2.330	3.060
<i>sppA</i>	Endopeptidase IV	2.167	4.131	1.161	2.933	1.008	1.815	0.025	0.053
<i>degP</i>	Periplasmic serine endoprotease DegP (	1.898	4.528	-0.654	1.204	2.194	3.385	-0.027	0.118
<i>dacA</i>	DD-peptidase	1.720	5.557	0.771	3.237	0.445	3.790	0.552	4.525
<i>pepB</i>	Peptidase B	1.320	4.344	1.243	2.978	2.210	3.727	1.005	4.160
<i>ptrA</i>	Protease III	1.276	3.248	0.644	1.356	0.258	1.795	-0.750	1.757
<i>pbgG</i>	Penicillin-binding protein 7	1.251	2.564	1.110	3.085	-1.969	3.807	-0.270	0.912
<i>degS</i>	Serine endoprotease DegS	-1.161	3.397	-0.620	1.589	0.390	1.121	-0.098	0.410
<i>iaaA</i>	Isoaspartyl peptidase	-1.908	3.744	-1.975	1.813	-1.491	3.991	-0.902	2.724
<i>iadA</i>	Isoaspartyl dipeptidase	-2.316	4.127	-2.092	2.171	-0.734	0.642	-1.374	2.891
<i>ompT</i>	Protease VII	-2.790	4.251	-1.403	2.132	-1.390	3.549	-1.062	1.689
<i>dacC</i>	Penicillin-binding protein 6	-0.382	3.786	-2.931	3.694	0.080	0.733	-1.396	3.784
<i>ldcA</i>	LD-carboxypeptidase A	-0.403	0.846	-1.729	1.918	-0.490	1.233	-1.345	2.672
<i>degQ</i>	Periplasmic serine endoprotease DegQ	-0.706	3.684	-1.590	2.468	0.237	2.191	-0.340	2.009
<i>pepT</i>	Peptidase T	-0.078	0.875	-1.321	2.217	-0.460	2.455	-1.423	4.477
<i>prlC</i>	Oligopeptidase A	1.103	3.144	0.092	0.550	-4.768	3.061	-5.470	3.838
<i>slt</i>	Exomuramidase	1.269	2.981	-0.947	1.543	1.146	3.258	-0.885	3.516
<i>srkA</i>	Stress response kinase A	1.266	3.134	-0.124	0.309	0.587	1.359	-0.477	2.445
<i>dsbA</i>	Thiol:disulfide interchange protein DsbA	1.705	3.244	0.353	0.683	-2.409	3.246	0.903	2.176
<i>ompC</i>	Outer membrane protein C	2.648	6.113	0.850	2.037	2.581	4.506	2.484	4.557
<i>yebE</i>	Inner membrane protein YebE	2.419	4.539	-0.220	0.996	3.702	4.940	0.114	0.138
<i>ygaU</i>	K <sup>+</sup> binding protein	1.820	2.839	-1.803	2.666	1.665	2.253	-2.703	4.043

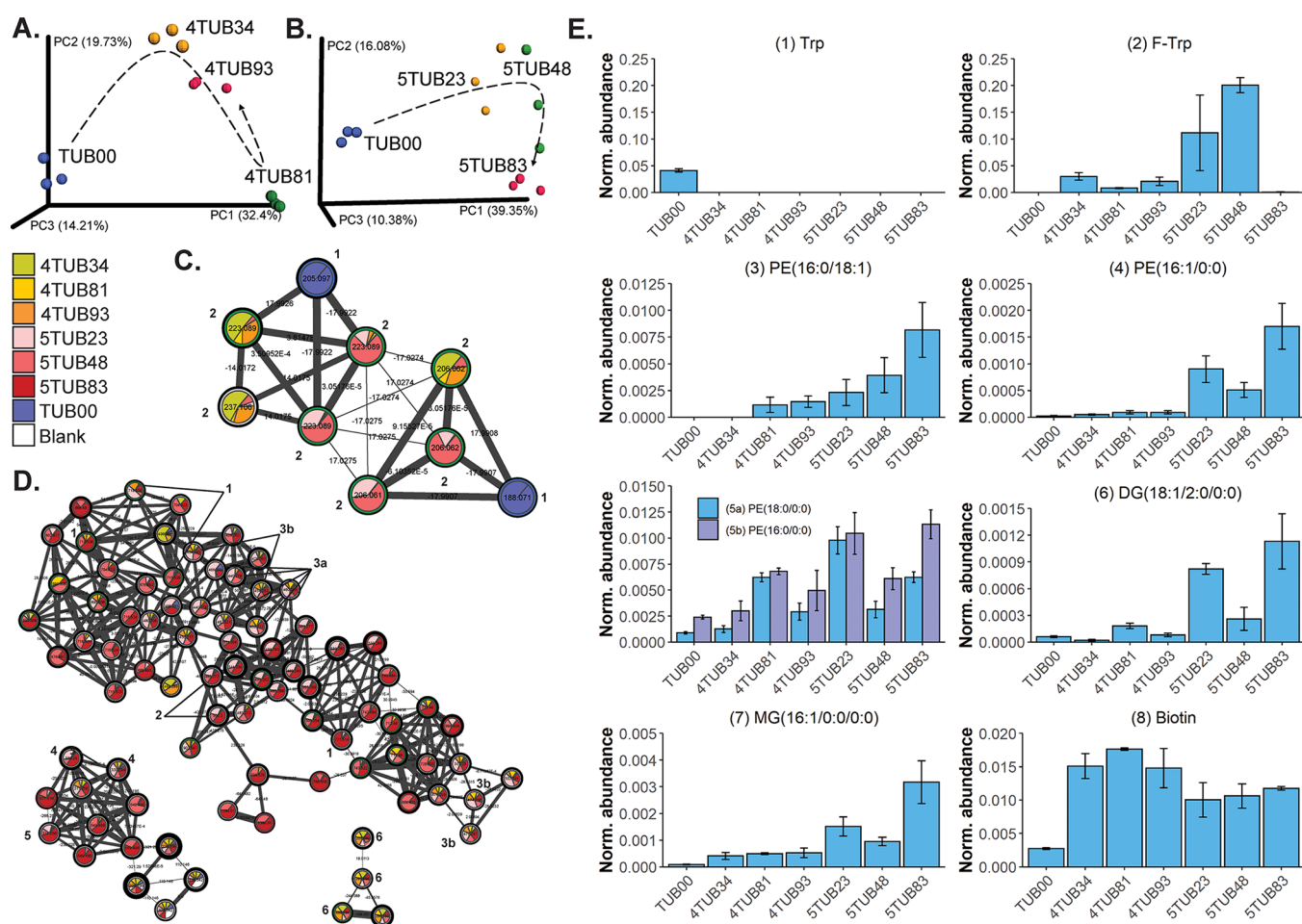
<sup>a</sup>The change in abundance is reported as log2-fold and is visualized in the form of a heat map, where green indicates upregulation (protein more abundant in the ALE isolates), and red indicates downregulation (protein more abundant in the TUB00 ancestor strain); white indicates a not significant abundance change. P-values are corrected by a permutation-based FDR.

proteomics Auxiliary Table 2), and the inactivation of this transporter in *E. coli* is associated with higher metabolic flux in the biosynthetic pathway of Ser,<sup>26</sup> which is required for the synthesis of (fluoro)tryptophan. Although the precise mechanism of indole uptake in the cells is still a matter of controversy, indole and its analogues are known to passively diffuse through cellular membranes.<sup>27</sup> Nonetheless, we found a mutation in *mtr*, encoding the high-affinity transporter Trp permease in all isolates from the 5-fluoroindole ALE experiment (Table 2). Since Trp is absent from the ALE medium, this mutation could inactivate an unused transporter (and thus be neutral) or facilitate 5-fluoroindole uptake into the cytoplasm and thus be beneficial for enlarging the pool of substrate for 5-fluorotryptophan synthesis. Proteomics data suggest that Mtr played a functional role in the adaptation toward fluoroindoles as xeno-nutrients, as it is upregulated throughout the whole 5-fluoroindole ALE as well as at the final time point of 4-fluoroindole ALE (4TUB93). Mutation of genes encoding multidrug efflux pumps (*mdtK*, *mdtO*, Table 2) was observed in the final isolate of 5-fluoroindole ALE (5TUB83) and at early stages of 4-fluoroindole ALE

(4TUB34, *mdtF*, Table 2). Multidrug efflux pumps carry out detoxification in the presence of xeno-compounds, such as the fluoroindoles.<sup>28,29</sup> However, in our ALE setup, precisely these compounds must be accumulated, as they are essential precursors for Trp and protein synthesis. For this reason, we believe that mutation of this class of transporters was crucial for the adaptation to fluoroindoles.

After genomics, we investigated the changes in proteome of *E. coli* upon global incorporation of fluorotryptophans. The fluorinated proteomes of the isolates from early and final time points of 4TUB and 5TUB were compared to the standard proteome of TUB00. Specifically, we quantified changes in abundance of single proteins by means of stable isotope labeling by amino acids in cell culture analysis (SILAC, Figures S10–S12).<sup>30</sup> Contrary to our expectation, relatively few proteins showed abundance change (see Figure S11). However, proteins and enzymes directly involved in the quality control of protein folding were strongly affected. In particular, protein chaperones and proteases that assist folding and degradation of misfolded proteins were upregulated at the early time points of both ALEs (green fields in columns





**Figure 4.** Analysis of metabolites produced during 4- and 5-fluoroindole ALE. Principal coordinate analysis (PCoA) plots with Canberra distance metric of the metabolomes from (A) 4-fluoroindole ALE and (B) 5-fluoroindole ALE. Each point represents the metabolome extracted from three independent cultures at early (yellow), intermediate (green), and final (red) time points of ALE and TUB00 (blue). The spatial distance in the plot is proportional to the chemical diversity between the samples, and evolutionary trajectories are shown (dashed arrows). Global Natural Products Social Molecular Networking (GNPS) visual output in the form of molecular subnetworks of (C) Trp and fluorotryptophan and of (D) lipids and biotin, i.e., the two groups of metabolites which were most pronounced among the isolates adapted to fluoroindoles. The nodes represent metabolites with unique retention time and  $m/z$  identifiers: (1) Trp; (2) fluorotryptophan; (3) 1-palmitoyl-2-oleyl-*sn*-glycero-3-phosphoethanolamine (PE(16:0/18:1)); (4) 1-palmitoleoyl-*sn*-glycero-3-phosphoethanolamine (PE(16:1/0:0)); (5a) 1-oleyl-2-hydroxy-*sn*-glycero-3-phosphoethanolamine (PE(18:0/0:0)); (5b) 1-palmitoyl-2-hydroxy-*sn*-glycero-3-phosphoethanolamine (PE(16:0/0:0)); (6) 1-oleoyl-2-acetyl-*sn*-glycerol (DG(18:1/2:0/0:0)); (7) monopalmitolein (MG(16:0/0:0/0:0)); and (8) biotin. The pie chart representation illustrates the relative abundance of each feature across the samples. (E) Normalized abundance of the annotated metabolites.

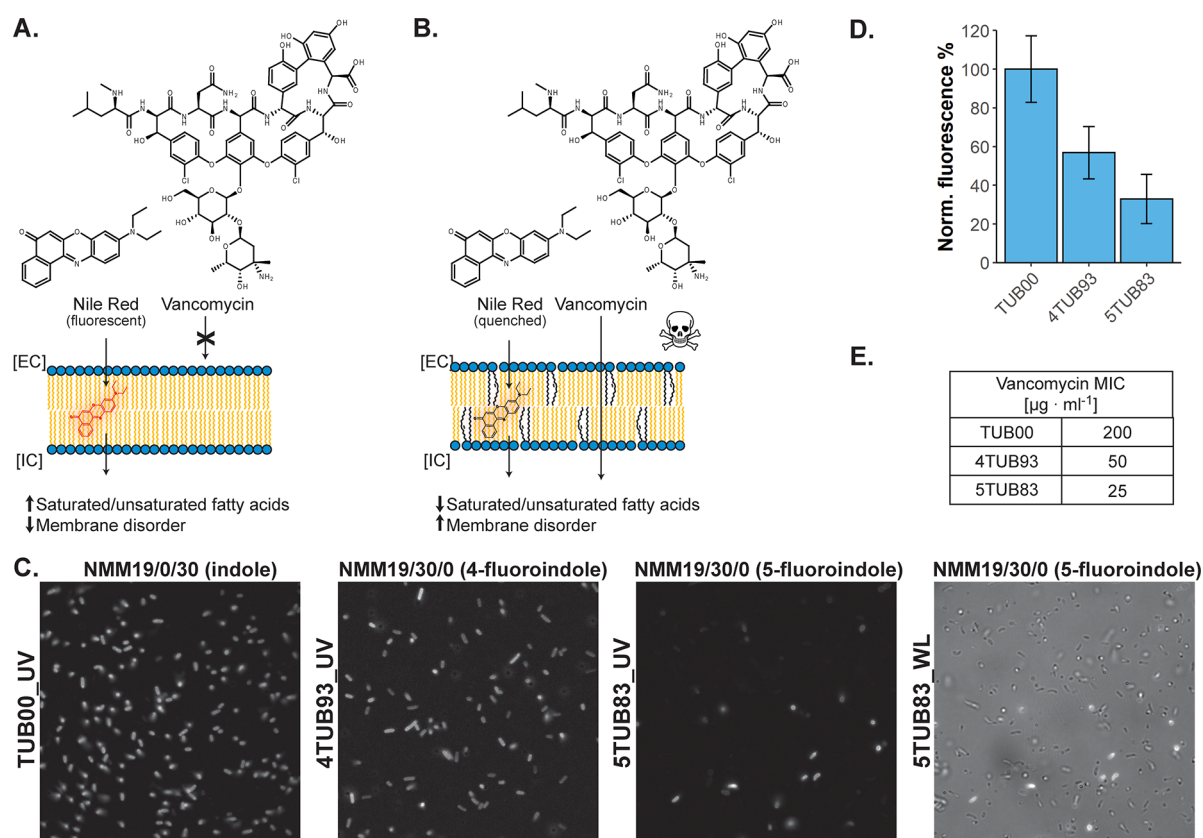
4TUB34 and 5TUB23, Table 3) and especially in 4TUB34, thus suggesting that a stress response associated with the presence of a large number of misfolded proteins was underway at the beginning of the adaptation process. This might be a specific effect of the incorporation of the fluorotryptophans into the proteome. Remarkably, at the end of ALE, the same proteins were downregulated back to ancestor strain levels (white fields in columns 4TUB93 and 5TUB83, Table 3), suggesting that the situation of stress was resolved.

We observed the differential expression of a subset of proteases and chaperones including DegP, PpiA, DsbA, and Spy, whose expression is known to be coregulated by the CpxAR system.<sup>31</sup> This system, composed of the sensor histidine kinase CpxA and the transcriptional regulator CpxR, controls the stress response against misfolded proteins in the periplasm of *E. coli*. Notably, at the end of both 4- and 5-fluoroindole ALE, DegP, PpiA, DsbA, Spy, and other CpxAR-regulated proteins such as Slt, YgaU,<sup>32</sup> SrkA,<sup>33</sup> OmpC,<sup>34</sup> and

YebE<sup>35</sup> were not upregulated, concomitant with a mutation of *cpxA* in the sequence encoding the histidine kinase domain required for cross-talking with CpxR (Table 2). We conclude that, although Trp is the least-abundant amino acid in the proteome of *E. coli* (~1%), its global fluorination induces a stress response, most likely associated with the partial misfolding of a large number of proteins. We hypothesize that cells coped with this condition by loosening the protein quality check mechanisms that normally ensure correct folding, such as protein chaperones and proteases.

Finally, we investigated whether the adaptation to fluoroindoles had altered the chemical composition of *E. coli*. The metabolomes of all relevant isolates were extracted and analyzed by nontargeted tandem-mass spectrometry and molecular networking with Global Natural Products Social Molecular Networking (GNPS, Auxiliary Table 3).<sup>36</sup> Multivariate statistics of principal coordinate analysis (PCoA) indicated that the metabolomes of TUB00 and the ALE isolates had increasingly diversified (Figure 4A,B). Moreover,





**Figure 5.** Cell membrane rearrangement during 4- and 5-fluoroindole ALE. (A, B) The hydrophobicity of *E. coli* membranes before and after the adaptation to fluoroindoles was probed by Nile Red staining, and the permeability properties were investigated by assessing the susceptibility to the antibiotic vancomycin. (C) Fluorescent micrographs of the ancestor strain TUB00 and of the final time points of 4- and 5-fluoroindole ALE (4TUB93 and 5TUB83, respectively) stained with Nile red. The last panel reports 5TUB83 irradiated with white light (WL) for cell count comparison. (D) Total fluorescence normalized to TUB00. (E) Minimal inhibitory concentration (MIC) of vancomycin.

investigation of GNPS molecular networks revealed that the most significant changes had occurred at the level of Trp, biotin, and lipid metabolites (Figure 4C,D and Table S11). In particular, Trp was present exclusively in TUB00, while fluorotryptophan was detected in the ALE isolates (Figure 4C,E and Table S11).

Remarkably, during the adaptation to 5-fluoroindole, (2) 5-fluorotryptophan became up to 5 times more abundant than Trp in TUB00 (Figure 4E, Table S11), which correlates with our hypothesis that mutation of TrpR upregulates 5-fluorotryptophan biosynthesis. [The isomers 4- and 5-fluorotryptophan cannot be distinguished by LC-MS/MS and GNPS. It was assumed that 4-fluorotryptophan was present in 4TUB and 5-fluorotryptophan in 5TUB, corresponding to the precursors supplied in the cultivation medium, 4- and 5-fluoroindole, respectively.] However, it might also indicate that this analogue accumulates in the cytoplasm due to inefficient turnover by TrpRS. The abundance of 5-fluorotryptophan abruptly dropped by  $10^3$ -fold at the end of ALE (5TUB83) concomitant with a mutation in *trpS*. The mutant TrpRS (Met187Leu, Table 2) might improve kinetics and increase the usage of 5-fluorotryptophan for protein biosynthesis, but this would require further investigation.

Besides the presence of fluorotryptophan, another distinctive trait of the fluoroindole-adapted strains was the presence of a large excess of biotin and glycerol- and phospholipids in comparison to the ancestor strain TUB00 (Figure 4D,E). Biotin is required for the first steps of fatty acid chain synthesis.

Indeed, at the final time point of 5-fluoroindole ALE (5TUB83), a high abundance of phosphatidylethanolamines (the main components of *E. coli* membranes) carrying unsaturated palmitic (C16) and oleic (C18) fatty acid chains was observed. It is known that indole naturally acts as a stressor for cells by acting as a membrane ionophore,<sup>20</sup> and this effect was expected to be stronger in the case of fluoroindoles due to their higher lipophilicity and polarity (see also Table S2).

For these reasons, we hypothesized that cell membrane rearrangement is required for the cells in the process of their specialization toward fluoroindoles. Enrichment of membranes with lipids carrying unsaturated fatty acids generally increases the fluidity of the lipid bilayer as it perturbs the stacking of adjacent chains and introduces disorder in the overall structure.<sup>37</sup> We investigated the properties of the cell membranes in 4TUB93, 5TUB83, and TUB00 by fluorescence microscopy. Cells were treated with Nile Red, a fluorescent dye often used to investigate membrane heterogeneity<sup>38</sup> and variation in composition of intracellular fatty acids<sup>39</sup> as well as to assess the hydrophobicity of the cellular membrane and wall of various microorganisms.<sup>40</sup> Nile Red is often regarded as a lipid-probe because it is highly fluorescent in nonpolar environments (e.g., the hydrophobic lipid bilayer of the cell membrane), but its fluorescence is quenched by an increase in polarity of the chemical environment or when exposed to polar solvents such as water<sup>41</sup> (Figure 5A,B). Both adapted strains were significantly less fluorescent than the ancestor strain

(Figure 5C,D) and particularly in the case of 5TUB83 where fluorescence is reduced by approximately 60% in comparison to TUB00 (Figure 5C,D), thus suggesting that the membrane properties had changed over the course of the adaptation to fluoroindoles.

We hypothesized that the rearrangement of the cell membrane might favor the uptake of fluoroindoles or reduce the toxic effect when these analogues accumulate in the lipid bilayer. Therefore, we tested the permeability of the cell membrane by measuring the susceptibility of TUB00, 4TUB93, and 5TUB83 to vancomycin, a high-molecular-weight antibiotic that is usually ineffective against *E. coli* as its passage through cell membranes is negligible (Figure 5A).<sup>42</sup> We found that both adapted strains are less tolerant toward vancomycin than TUB00 (Figure 5E, Table S12). This confirmed that adaptation toward fluoroindoles affected the organization of the cell membrane and increased its permeability to extracellular solutes.

## DISCUSSION

Here, we report the first laboratory adaptation of *E. coli* toward the usage of two different fluorinated indole analogues as precursors for endogenous protein synthesis that eventually led to the complete incorporation of 4- and 5-fluorotryptophan into the proteome of actively proliferating cells. An integrative multiomics analysis of the adaptation process demonstrated that ALE selected for strains with reconfigured regulatory networks, albeit carrying surprisingly few genetic alterations. Evolutionary mechanisms originating from approximately 30 mutations were accompanied by only minor changes in the protein synthesis machinery. However, adaptation deeply affected core processes such as protein folding, membrane dynamics, stress, and nutrient uptake.

Based on our results, we believe that at the initial ALE stages, the key features of adaptation included upregulation of proteases and protein chaperones as well as activation of the CpxAR-mediated stress response. As cells became proficient in utilizing fluoroindoles, these processes were attenuated. This change in expression can be explained based on the assumption that at the beginning of the adaptation a high number of newly fluorinated proteins are misfolded and trigger the activation of the stress response. Under these conditions, cells invest a considerable fraction of metabolic energy in the synthesis of proteases and chaperones. However, since the misfolding issue cannot be resolved, and fluorinated proteins keep on being synthesized, a toxic, energy-dissipating cycle is established. At this stage, those cell subpopulations that acquired mutations that reduce the overexpression of proteases and chaperones (i.e., that attenuate the misfolded protein stress response) have a fitness advantage. This is strong evidence that the interplay between protein mutational robustness, protein folding, and environmental stress is a key factor that determines the evolution of new traits in habitats containing fluorinated amino acids. The challenge of continuous exposure to fluorinated stressors was responded to primarily at the level of the cell membrane, as demonstrated by evidence of membrane rearrangement and diversification of lipid production, especially during adaptation to 5-fluoroindole. The change in lipid composition does not compromise the viability of the strains and reflects the difference in lipophilicity and polarity between the two fluorinated indole analogues. Upregulated production of unsaturated phosphatidylethanolamines changes membrane properties and might

contribute to rendering the adapted strains more susceptible toward the antibiotic vancomycin.

Furthermore, it is reasonable that adaptation is facilitated by mutations in transporter proteins such as Mtr and multidrug efflux pumps (MdtF, MdtO, MdtK) that favor intracellular accumulation of the fluorinated xenobiotic compounds. 4TUB and 5TUB strains developed the ability to grow in a fluorinated habitat under conditions that are not permissive to the parent strain TUB00 (Figure S14). Their growth behavior suffered by significant elongation of the generation time when moving from complete media containing amino acids (e.g., NMM19) to media containing only fluoroindoles and no amino acids (NMM0, Table S13). During ALE not only indole was depleted from the cultivation medium, but also amino acids. This eliminated any possible source of contamination by Trp, but at the same time, it came at a high metabolic cost as it required the cells to biosynthesize the missing amino acids in order to carry on protein synthesis. However, it must be noted that the generation time of the adapted strains in the medium containing all amino acids and only indole is longer than that of TUB00 (NMM19/0/30:189 min of 4TUB93 and 171 min of 5TUB83, i.e., 1.5 and 1.4 times in comparison to 126 min of TUB00, respectively; Table S13), which hints that indole/Trp do not contribute to the fitness of the strains after ALE. On the other hand, the adapted strains show shorter generation times in media containing the fluorinated indole analogues (30:1 concentration ratio to indole, all amino acids provided, NMM19/30/1). When 5-fluoroindole is supplemented, 5TUB83 grows 1.2-times faster than TUB00 (145 min/generation TUB00; 115 min/generation 5TUB83; Table S13) and 1.5-times faster than in medium containing only indole (NMM19/0/30, 171 min/generation; Table S14). When 4-fluoroindole is supplemented, the difference between 4TUB93 and TUB00 is smaller (163 min TUB00; 158 min 4TUB93; Table S13) which reflects the physicochemical similarity of this analogue to indole/Trp as shown in Figure S1 ("Structure, size, lipophilicity, and dipole moment of indole and fluoroindoles" in the Supporting Information). However, 4TUB93 grows 1.2 times faster in the presence of 4-fluoroindole rather than indole (NMM19/0/30, 189 min/generation; Table S13). Taken together, we believe that these data provide strong evidence for the adaptation of 4TUB and 5TUB to the fluorinated indole and Trp analogues.

The multiomics data collected on the adapted strains 4TUB and 5TUB provided us with a solid basis for understanding the effect of fluorinated indole/Trp analogues on cellular metabolism, which is fundamental for application in synthetic biology and biotechnology (i.e., optimizing the synthesis of protein variants and remove sources of metabolic stress). Further studies are needed to identify the core biological barriers that control microbial adaptation to unnatural chemistries and engineering microbial strains with improved growth behavior for synthetic scopes, e.g., facile fluorine-labeling of peptides and whole proteins. To date, the incorporation efficiency of fluorinated amino acids via the classical molecular biological methods (e.g., amber stop codon suppression and sense codon-reassignment) is far from 100%. The occasional translation of codons by the corresponding canonical amino acid results in a heterogeneous mixture of native and partially labeled proteins, which are impossible to separate by chromatography. According to the approach we show here, removing the dependency of host cells on the supplementation of canonical amino acids for growth allows

one to synthesize only fully labeled proteins and to reproducibly obtain homogeneous products. This is particularly appealing for the synthesis of bioactive proteins and drug peptides for clinical applications.

In future studies, we envision expanding ALE coupled with multiomics analysis to a broad range of fluorinated amino acid analogues. Such large data sets could be then assembled into normalized compendia to provide an ideal training ground for modeling algorithms to predict evolutionary landscapes. Eventually, this will enable us to define one comprehensible metabolic model of adaptation and to remove the bottlenecks that limit life based on unnatural amino acids.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acscentsci.0c00679>.

Auxiliary Table 1 (XLSX)

Auxiliary Table 2 (XLSX)

Auxiliary Table 3 (XLSX)

Additional experimental details, data, and figures including structures, *E. coli* genotype, *trp* operon amplification, GC-MS results, OD<sub>600</sub> and incubation time, SILAC experimental setup, volcano plots, histograms, and growth curves (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Authors

**Nediljko Budisa** – Institute of Chemistry & Biocatalysis, Technische Universität Berlin, Berlin 10623, Germany; Department of Chemistry, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada; [orcid.org/0000-0001-8437-7304](https://orcid.org/0000-0001-8437-7304); Email: [nediljko.budisa@tu-berlin.de](mailto:nediljko.budisa@tu-berlin.de)

**Beate Koksche** – Institute of Chemistry and Biochemistry, Organic Chemistry, Freie Universität Berlin, Berlin 14195, Germany; [orcid.org/0000-0002-9747-0740](https://orcid.org/0000-0002-9747-0740); Email: [beate.koksche@fu-berlin.de](mailto:beate.koksche@fu-berlin.de)

### Authors

**Federica Agostini** – Institute of Biochemistry & Charité Universitätsmedizin Berlin, Berlin 10117, Germany

**Ludwig Sinn** – Institute of Biotechnology & Bioanalytics, Technische Universität Berlin, Berlin 10623, Germany

**Daniel Petras** – Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, California 92093, United States; [orcid.org/0000-0002-6561-3022](https://orcid.org/0000-0002-6561-3022)

**Christian J. Schipp** – Institute of Chemistry & Biocatalysis, Technische Universität Berlin, Berlin 10623, Germany; [orcid.org/0000-0002-6271-969X](https://orcid.org/0000-0002-6271-969X)

**Vladimir Kubyshkin** – Institute of Chemistry & Biocatalysis, Technische Universität Berlin, Berlin 10623, Germany; Department of Chemistry, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada

**Allison Ann Berger** – Institute of Chemistry and Biochemistry & Organic Chemistry, Freie Universität Berlin, Berlin 14195, Germany

**Pieter C. Dorrestein** – Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, California 92093, United States

**Juri Rappsilber** – Institute of Biotechnology & Bioanalytics, Technische Universität Berlin, Berlin 10623, Germany;

Wellcome Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3BF, United Kingdom; [orcid.org/0000-0001-5999-1310](https://orcid.org/0000-0001-5999-1310)

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acscentsci.0c00679>

## Author Contributions

F.A. designed the experimental setup, performed strain cultivation and all other experiments and wrote the manuscript. D.P. performed metabolomics MS data acquisition, supervised molecular networking data analysis and assisted in writing the manuscript. L.S. performed shotgun proteomics data acquisition, SILAC raw data analysis and assisted in writing the manuscript. C.J.S. performed the measurements of ATP/[<sup>32</sup>P]pyrophosphate exchange assay. V.K. performed the measurements of fluorindoles dipole moment and lipophilicity. A.A.B. assisted in designing experiments and in writing the manuscript. P.C.D. directed metabolomics experiments. J.R. directed proteomics experiments and assisted in writing the manuscript. N.B. conceived the project, directed experimental work, and wrote the manuscript. B.K. designed the experiments and wrote the manuscript.

## Funding

This work was supported by the US National Science Foundation (NSF) Inspire Track II IOS-1343020, National Institutes of Health (NIH) Grants GMS10RR029121, SP41GM103484-07, the Deutsche Forschungsgemeinschaft (DFG) with Grant PE 2600/1, RA 2365/4-1, and 25065445, the Einstein Foundation and the Wellcome Trust through a Senior Research Fellowship to J.R. [103139] and a multiuser equipment grant [108504]. The Wellcome Centre for Cell Biology is supported by core funding from the Wellcome Trust [203149].

## Notes

The authors declare no competing financial interest.

Data and Materials Availability: The complete genomic, proteomic, and metabolomic data sets are available as auxiliary tables. All metabolomics MS/MS data can be found on the Mass spectrometry Interactive Virtual Environment (MassIVE) at <https://massive.ucsd.edu/> with the accession number MSV000083134. Molecular Networking and Spectrum Library Matching results can be found online at GNPS under the following links: <https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=628bd0896fad4eae91f6be5f12373d25> (analogue search on); <https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=7b23638cf23a4cb5ab71abbbd6a01248> (analogue search off). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository<sup>43</sup> with the dataset identifier PXD011771.

## ■ ACKNOWLEDGMENTS

The authors thank Dr. Franz-Josef Schmitt (TU Berlin) for cell imaging by fluorescence microscopy and Dr. Torsten Semmler (Robert Koch Institut, Berlin) for revision of sequencing data.

## ■ REFERENCES

(1) Harper, D. B.; O'Hagan, D.; Murphy, C. D. Fluorinated Natural Products: Occurrence and Biosynthesis. *Handbook of Environmental Chemistry* **2003**, 3, 141–169.



- (2) O'Hagan, D.; Deng, H. Enzymatic Fluorination and Biotechnological Developments of the Fluorinase. *Chem. Rev.* **2015**, *115* (2), 634–649.
- (3) Müller, K.; Faeh, C.; Diederich, F. Fluorine in Pharmaceuticals: Looking beyond Intuition. *Science* **2007**, *317* (5846), 1881–1886.
- (4) O'Hagan, D. Understanding Organofluorine Chemistry. An Introduction to the C-F Bond. *Chem. Soc. Rev.* **2008**, *37* (2), 308–319.
- (5) Berger, R.; Resnati, G.; Metrangolo, P.; Weber, E.; Hulliger, J. Organic Fluorine Compounds: A Great Opportunity for Enhanced Materials Properties. *Chem. Soc. Rev.* **2011**, *40* (7), 3496–3508.
- (6) Jackson, J. T.; Hammill, A.; Mehl, R. A. Site-Specific Incorporation of a 19F-Amino Acid into Proteins as an NMR Probe for Characterizing Protein Structure and Reactivity. *J. Am. Chem. Soc.* **2007**, *129* (5), 1160–1166.
- (7) Salwiczek, M.; Nyakatura, E. K.; Gerling, U. I. M.; Ye, S.; Koks, B. Fluorinated Amino Acids: Compatibility with Native Protein Structures and Effects on Protein–Protein Interactions. *Chem. Soc. Rev.* **2012**, *41* (6), 2135–2171.
- (8) Berger, A. A.; Völler, J.-S.; Budisa, N.; Koks, B. Deciphering the Fluorine Code—The Many Hats Fluorine Wears in a Protein Environment. *Acc. Chem. Res.* **2017**, *50* (9), 2093–2103.
- (9) Stockholm Convention. Thirteenth meeting of the Persistent Organic Pollutants Review Committee (POPRC.13).
- (10) Browne, D. T.; Kenyon, G. L.; Hegeman, G. D. Incorporation of Monofluorotryptophans into Protein during the Growth of *Escherichia Coli*. *Biochem. Biophys. Res. Commun.* **1970**, *39* (1), 13–19.
- (11) Pratt, E. A.; Chien, H. Incorporation of Fluorotryptophans into Proteins of *Escherichia Coli*. *Biochemistry* **1975**, *14* (13), 3035–3040.
- (12) Richmond, M. H. The Effect of Amino Acid Analogues on Growth and Protein Synthesis in Microorganisms. *Bacteriol. Rev.* **1962**, *26*, 398–420.
- (13) Wong, J. T. Membership Mutation of the Genetic Code: Loss of Fitness by Tryptophan. *Proc. Natl. Acad. Sci. U. S. A.* **1983**, *80* (20), 6303–6306.
- (14) Bacher, J. M.; Ellington, A. D. Selection and Characterization of *Escherichia Coli* Variants Capable of Growth on an Otherwise Toxic Tryptophan Analogue. *J. Bacteriol.* **2001**, *183* (18), 5414–5425.
- (15) Rennert, O. M.; Anker, H. S. On the Incorporation of 5'S',5'-Trifluoroisoleucine into Proteins of *E. Coli*. *Biochemistry* **1963**, *2* (3), 471–476.
- (16) Mat, W.-K.; Xue, H.; Wong, J. T.-F. Genetic Code Mutations: The Breaking of a Three Billion Year Invariance. *PLoS One* **2010**, *5* (8), e12206–e12206.
- (17) Yu, A. C. S.; Yim, A. K. Y.; Mat, W. K.; Tong, A. H. Y.; Lok, S.; Xue, H.; Tsui, S. K. W.; Wong, J. T. F.; Chan, T. F. Mutations Enabling Displacement of Tryptophan by 4-Fluorotryptophan as a Canonical Amino Acid of the Genetic Code. *Genome Biol. Evol.* **2014**, *6* (3), 629–641.
- (18) Bacher, J. M.; Hughes, R. A.; Tze-Fei Wong, J.; Ellington, A. D. Evolving New Genetic Codes. *Trends Ecol. Evol.* **2004**, *19* (2), 69–75.
- (19) Hoesl, M. G.; Oehm, S.; Durkin, P.; Darmon, E.; Peil, L.; Aerni, H.-R.; Rappsilber, J.; Rinehart, J.; Leach, D.; Söll, D.; Budisa, N. Chemical Evolution of a Bacterial Proteome. *Angew. Chem., Int. Ed.* **2015**, *54* (34), 10030–10034.
- (20) Chimere, C.; Murray, A. J.; Oldewurtel, E. R.; Summers, D. K.; Keyser, U. F. The Effect of Bacterial Signal Indole on the Electrical Properties of Lipid Membranes. *ChemPhysChem* **2013**, *14* (2), 417–423.
- (21) Choi, Y.; Sims, G. E.; Murphy, S.; Miller, J. R.; Chan, A. P. Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLoS One* **2012**, *7* (10), e46688.
- (22) Choi, Y.; Chan, A. P. PROVEAN Web Server: A Tool to Predict the Functional Effect of Amino Acid Substitutions and Indels. *Bioinformatics* **2015**, *31* (16), 2745–2747.
- (23) Budisa, N. Prolegomena to Future Experimental Efforts on Genetic Code Engineering by Expanding Its Amino Acid Repertoire. *Angew. Chem., Int. Ed.* **2004**, *43* (47), 6426–6463.
- (24) Choe, D.; Lee, J. H.; Yoo, M.; Hwang, S.; Sung, B. H.; Cho, S.; Palsson, B.; Kim, S. C.; Cho, B.-K. Adaptive Laboratory Evolution of a Genome-Reduced *Escherichia Coli*. *Nat. Commun.* **2019**, *10* (1), 935.
- (25) Marmorstein, R. Q.; Sigler, P. B. Stereochemistry Effects of L-Tryptophan and Its Analogues on Trp Repressor's Affinity for Operator-DNA. *J. Biol. Chem.* **1989**, *264* (16), 9149–9154.
- (26) Lu, N.; Zhang, B.; Cheng, L.; Wang, J.; Zhang, S.; Fu, S.; Xiao, Y.; Liu, H. Gene Modification of *Escherichia Coli* and Incorporation of Process Control to Decrease Acetate Accumulation and Increase L-Tryptophan Production. *Ann. Microbiol.* **2017**, *67* (8), 567–576.
- (27) Piñero-Fernandez, S.; Chimere, C.; Keyser, U. F.; Summers, D. K. Indole Transport across *Escherichia Coli* Membranes. *J. Bacteriol.* **2011**, *193* (8), 1793–1798.
- (28) Bohnert, J. A.; Schuster, S.; Fährnich, E.; Trittler, R.; Kern, W. V. Altered Spectrum of Multidrug Resistance Associated with a Single Point Mutation in the *Escherichia Coli* RND-Type MDR Efflux Pump YhiV (MdtF). *J. Antimicrob. Chemother.* **2007**, *59* (6), 1216–1222.
- (29) Pietsch, F.; Bergman, J. M.; Brandis, G.; Marcusson, L. L.; Zorzet, A.; Huseby, D. L.; Hughes, D. Ciprofloxacin Selects for RNA Polymerase Mutations with Pleiotropic Antibiotic Resistance Effects. *J. Antimicrob. Chemother.* **2017**, *72* (1), 75–84.
- (30) Ong, S.-E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics. *Mol. Cell. Proteomics* **2002**, *1* (5), 376–386.
- (31) Pogliano, J.; Lynch, A. S.; Belin, D.; Lin, E. C.; Beckwith, J. Regulation of *Escherichia Coli* Cell Envelope Proteins Involved in Protein Folding and Degradation by the Cpx Two-Component System. *Genes Dev.* **1997**, *11* (9), 1169–1182.
- (32) Bernal-Cabas, M.; Ayala, J. A.; Raivio, T. L. The Cpx Envelope Stress Response Modifies Peptidoglycan Cross-Linking via the L,D-Transpeptidase LdtD and the Novel Protein YgaU. *J. Bacteriol.* **2015**, *197* (3), 603–614.
- (33) Dorsey-Oresto, A.; Lu, T.; Mosel, M.; Wang, X.; Salz, T.; Drlica, K.; Zhao, X. YihE Kinase Is a Central Regulator of Programmed Cell Death in Bacteria. *Cell Rep.* **2013**, *3* (2), 528–537.
- (34) Batchelor, E.; Walther, D.; Kenney, L. J.; Goulian, M. The *Escherichia Coli* CpxA-CpxR Envelope Stress Response System Regulates Expression of the Porins OmpF and OmpC. *J. Bacteriol.* **2005**, *187* (16), 5723–5731.
- (35) Yamamoto, K.; Ishihama, A. Characterization of Copper-Inducible Promoters Regulated by CpxA/CpxR in *Escherichia Coli*. *Biosci., Biotechnol., Biochem.* **2006**, *70* (7), 1688–1695.
- (36) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W. T.; Crüsemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderón, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C. C.; Floros, D. J.; Gavilan, R. G.; Kleigrew, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C. C.; Yang, Y. L.; Humpf, H. U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; Boya, C. A. P.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodriguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P. M.; Phapale, P.; Nothias, L. F.; Alexandrov, T.; Litaudon, M.; Wolfender, J. L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D. T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Müller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B.; Pogliano, K.; Lington, R. G.; Gutiérrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N. Sharing



and Community Curation of Mass Spectrometry Data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34* (8), 828–837.

(37) Zhang, Y.-M.; Rock, C. O. Membrane Lipid Homeostasis in Bacteria. *Nat. Rev. Microbiol.* **2008**, *6*, 222–233.

(38) Ira; Krishnamoorthy, G. Probing the Link between Proton Transport and Water Content in Lipid Membranes. *J. Phys. Chem. B* **2001**, *105* (7), 1484–1488.

(39) Guzmán, H. M.; de la Jara Valido, A.; Duarte, L. C.; Presmanes, K. F. Estimate by Means of Flow Cytometry of Variation in Composition of Fatty Acids from *Tetraselmis Suecica* in Response to Culture Conditions. *Aquacult. Int.* **2010**, *18* (2), 189–199.

(40) Thrane, C.; Olsson, S.; Harder Nielsen, T.; Sorensen, J. Vital Fluorescent Stains for Detection of Stress in *Pythium Ultimum* and *Rhizoctonia Solani* Challenged with Viscosinamide from *Pseudomonas Fluorescens* DR54. *FEMS Microbiol. Ecol.* **1999**, *30* (1), 11–23.

(41) Greenspan, P.; Fowler, S. D. Spectrofluorometric Studies of the Lipid Probe, Nile Red. *J. Lipid Res.* **1985**, *26* (7), 781–789.

(42) Huang, K. C.; Mukhopadhyay, R.; Wen, B.; Gitai, Z.; Wingreen, N. S. Cell Shape and Cell-Wall Organization in Gram-Negative Bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (49), 19282–19287.

(43) Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D. J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M.; Pérez, E.; Uszkoreit, J.; Pfeuffer, J.; Sachsenberg, T.; Yilmaz, S.; Tiwary, S.; Cox, J.; Audain, E.; Walzer, M.; Jarnuczak, A. F.; Ternent, T.; Brazma, A.; Vizcaino, J. A. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **2019**, *47*, D442–D450.